Robustness and device independence of verifiable blind quantum computing

Robustness and device independence of verifiable blind quantum computing

2015 New J. Phys. 17 083040

(http://iopscience.iop.org/1367-2630/17/8/083040)

New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG
IOP Institute of Physics

Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

OPEN ACCESS

RECEIVED
15 April 2015

REVISED
16 June 2015

ACCEPTED FOR PUBLICATION
10 July 2015

PUBLISHED
19 August 2015

PAPER

Robustness and device independence of verifiable blind quantum computing

Alexandru Gheorghiu[1], **Elham Kashefi**[1,2] **and Petros Wallden**[1]

[1] School of Informatics, University of Edinburgh, 10 Crichton Street, Edinburgh EH8 9AB, UK
[2] CNRS LTCI, Departement Informatique et Reseaux, Telecom ParisTech, Paris CEDEX 13, France

E-mail: a.gheorghiu@sms.ed.ac.uk

Keywords: delegated quantum computation, quantum verification, device independence, composition, fault tolerance

Abstract

Recent advances in theoretical and experimental quantum computing bring us closer to scalable quantum computing devices. This makes the need for protocols that verify the correct functionality of quantum operations timely and has led to the field of quantum verification. In this paper we address key challenges to make quantum verification protocols applicable to experimental implementations. We prove the robustness of the single server verifiable universal blind quantum computing protocol of Fitzsimons and Kashefi (2012 arXiv:1203.5217) in the most general scenario. This includes the case where the purification of the deviated input state is in the hands of an adversarial server. The proved robustness property allows the composition of this protocol with a device-independent state tomography protocol that we give, which is based on the rigidity of CHSH games as proposed by Reichardt *et al* (2013 *Nature* **496** 456–60). The resulting composite protocol has lower round complexity for the verification of entangled quantum servers with a classical verifier and, as we show, can be made fault tolerant.

1. Introduction

While the prospect of commercially available universal quantum computing is still distant, a number of experiments involving multi-qubit systems have recently been developed. Irrespective of their applications, these technologies require methods and tools for verifying the correctness of their operations. Assuming that quantum computing is more powerful than classical computing, a simulation-based approach for quantum verification of devices with sufficiently large number of qubits, becomes practically impossible. Aaronson and Arkhipov showed in [3] that even a rudimentary quantum computer constructed with linear-optical elements cannot be efficiently simulated. Similarly, verifying the correct preparation of a general n qubit state via state tomography also involves exponential overhead since it requires collecting statistics from 4^n separate observables [4].

The verification of quantum devices becomes more complicated when the functionality involves cryptographic primitives. In these cases, incorrect operations could be the result of actions of an adversary. Thus it becomes necessary to guarantee the security of the application under certain assumptions about the devices. Ideally a protocol should remain secure even if the devices are faulty and partially controlled by adversaries. This would lead to a solution that is *device-independent* and *robust*. However, generating such protocols has proven difficult. Even in quantum key distribution, a complete proof of security for a device-independent protocol, in the presence of noise, has been achieved only recently [5].

The issue of verification needs to be resolved to be able to exploit successfully any future quantum computers. Moreover, one expects that the first large scale quantum devices are unlikely to be personal computers. Instead, they will probably function as *servers* to which *clients* can connect and request the computation of some difficult problem. The client may also require his computation to be private, i.e. require that the server does not learn anything about it. We should therefore construct protocols that verify an arbitrary delegated quantum computation and prove the security and correctness of this verification technique.

The approaches that have been so far successful are those based on *interactive proof systems* [6, 7], where a *trusted*, computationally limited verifier (also known as client, in a cryptographic setting) exchanges messages with an *untrusted*, powerful quantum prover, or multiple provers (also known as servers). The verifier attempts to certify that, with high probability, the provers are performing the correct quantum operations. Because we are dealing with a new form of computation, the verification protocols, while based on established techniques are fundamentally different from their classical counterparts. A number of quantum verification protocols have been developed, for different functionalities of devices and using a variety of different strategies to achieve verification [1, 2, 8–17]. The assumptions made depend on the specific target and desired properties of the protocol. For example, if the emphasis is on creating an immediate practical implementation, then this should be reflected in the technological requirements leading to a testable application with current technology [17]. Alternatively, if the motivation is to prove a theoretical result, we may relax some requirements such as efficient scaling [2]. An important open problem in the field of quantum verification, is whether a scheme with a fully classical verifier is possible [18, 19]. We know, however, that verification *is* possible in the following two scenarios.

(1) A verifier with minimal quantum capacity (ability to prepare random single qubits) and a single quantum prover [1]. This is the Fitzsimons and Kashefi (FK) protocol.

(2) A fully classical verifier and two non-communicating quantum provers that share entanglement [20]. This is the Reichardt, Unger and Vazirani (RUV) protocol.

One of our objectives is to obtain a *device-independent* (allowing untrusted quantum devices) version of the FK protocol, by composing it with the RUV protocol.

Here we should make some remarks as to what constitutes a device-independent protocol. A device-independent protocol is a protocol where the honest participants do not trust their devices. They obtain only classical outcomes and they assume the worse, i.e. that those devices were prepared by adversaries and were programmed to function in the most adversarial way. This means they could be exploiting pre-shared randomness with other devices and adversaries, in such a way that any correlated attack is possible. In this view, the verification protocols of [1] and [21, 22] are not device-independent because the honest client has to assume that the output of his device is *not* correlated with the device or actions of the malicious prover. In our work, we assume a fully classsical client, that uses two malicious but non-communicating provers. One of the provers is only required to perform measurements and so can be viewed as an untrusted measurement device, not necessarily a full quantum computer. In the independent work of [23], a device independent verification protocol with a single prover and a client having an untrusted device, is proposed. These two settings are two different views of the same situation. Since the device of the client is not trusted and is a black-box, it can be modelled as a prover, that uses pre-shared randomness with the other prover to deceive the client.

Additional properties we aim to achieve from the composition of the FK and RUV protocol are *fault tolerance* (allowing noisy devices) and *reduced round complexity*. Composing protocols can indeed be fruitful since it could lead to new protocols that inherit the advantages of both constituents. The universal composabillity framework, allowing for secure compositions, has been successfully extended to the quantum realm [24–26]. Recently, the security of single server verifiable universal blind quantum computing protocols has been demonstrated in an abstract cryptographic framework [27] that is also known to be equivalent to the simulation-based composability framework. However this setting does not fulfil the necessary requirements for our composition. This is because, when combining a single server verification scheme (FK) with an entangled server scheme (RUV), there exists the possibility of correlated attacks, which are not explicitly treated in the composability framework. Such attacks can occur when an untrusted server's strategy is correlated with deviations in the protocol's input state. Our robustness result resolves this issue in the stand alone composition setting and the same technique could potentially be extended to the composable framework of [27], thus resolving the problem of correlated attacks.

The type of composition that we require is sequential. We take the output of the first protocol and use it as input for the second. However, in general, the output of the first protocol is not necessarily an acceptable input for the second protocol. In particular, since the verification protocols are probabilistic, their outputs typically deviate by a small amount from the ideal one. Thus, it is necessary that the second protocol remain secure even if the input is slightly deviated from the ideal one. Moreover, we make sure that adversaries cannot exploit any correlations between the deviated input and their strategy to compromise the security of the protocol. Therefore to securely compose the protocols, we need to address these new type of attacks. The main results of this paper can be summarized as follows.

(1) We prove that the FK protocol is strongly robust, see theorem 1. First, we show that FK can tolerate inputs which deviate from their ideal values by a small amount, see lemma 7. However, for composition with other protocols a stronger property is needed. We therefore proceed to show that the FK protocol is robust even when the deviated input is correlated with an external system possessed by an adversary (for example the provers' private systems in the RUV protocol), see lemma 8.

(2) An immediate consequence of the robustness theorem is that we can construct a composite protocol combining RUV with FK. The required input states for the FK protocol are prepared via the state tomography sub-protocol of RUV. Our composite protocol inherits the device independence property of RUV, see theorem 3. Additionally, since we do not require the full RUV protocol, the composite protocol also has an improved round complexity. Moreover, since one of the provers is only required to run the state tomography sub-protocol, it only needs to perform measurements. This means that only one of the provers should be considered a quantum computer (the one running the quantum computation), whereas the other can be viewed as an untrusted quantum measurement device.

(3) Lastly, we address the distinction between robustness and fault tolerance and show how the FK protocol can be made fault tolerant, thereby making our proposed composite approach fault tolerant as well. Here we should note that in [1], fault-tolerant computation and quantum error correcting codes (QEC) were used, but with specific aim to boost the security of the protocol from polynomial to exponential. In particular, the traps were *not* encoded using QEC, and thus the protocol was not tolerant to error-prone devices, since single faults on trap qubits would lead to an honest run being aborted. As we will see the FK protocol can be made fault-tolerant, but more care is needed.

In section 1.1 we give some preliminaries. In section 2 we present the main results, that we summarized above, and outline their proofs. In particular we give robustness in section 2.1, composition in section 2.2 and fault tolerance in section 2.3. Further details of the proofs are given in section 3 for robustness, in section 4 for composition and in section 5 for fault tolerance. We conclude in section 6.

1.1. Preliminaries

We first introduce the relevant concepts used in describing verification protocols and then briefly present the two protocols we will built on (FK and RUV).

1.1.1. Interactive proof systems

As explained in [6, 10], a language $\mathcal{L}$ is said to admit an interactive proof system if there exists a computationally unbounded prover $\mathcal{P}$ and a BPP verifier $\mathcal{V}$, such that for any $x \in \mathcal{L}$, $\mathcal{P}$ convinces $\mathcal{V}$ that $x \in \mathcal{L}$ with probability $\geqslant \frac{2}{3}$. Additionally, when $x \notin \mathcal{L}$, $\mathcal{P}$ convinces $\mathcal{V}$ that $x \in \mathcal{L}$ with probability $\leqslant \frac{1}{3}$. Mathematically, we have the following two conditions[3].

- Completeness: $Pr\,(\mathcal{V} \leftrightarrow \mathcal{P} \text{ accepts } x \;\mid\; x \in \mathcal{L}) \geqslant \frac{2}{3}$.
- Soundness: $Pr\,(\mathcal{V} \leftrightarrow \mathcal{P} \text{ accepts } x \;\mid\; x \notin \mathcal{L}) \leqslant \frac{1}{3}$.

The set of languages which admit such an interactive proof system define the complexity class IP. We are interested in the case when the prover is a polynomial-time quantum computer (i.e. a BQP machine). In [10], the first definition of such a quantum interactive proof system was given, which we use here:

Definition 1 [10]. Quantum prover interactive proof (QPIP) is an interactive proof system with the following properties.

(i) The prover is computationally restricted to BQP.

(ii) The verifier is a hybrid quantum-classical machine. Its classical part is a BPP machine. The quantum part is a register of c qubits (for some constant c), on which the prover can perform arbitrary quantum operations. At any given time, the verifier is not allowed to possess more than c qubits. The interaction between the quantum and classical parts is the usual one: the classical part controls which operations are to be performed on the quantum register, and outcomes of measurements of the quantum register can be used as input to the classical machine.

[3] Note that completeness can be viewed as the probability of the verifier accepting when the prover is honest. Similarly, soundness is the probability of accepting, when the prover is dishonest.

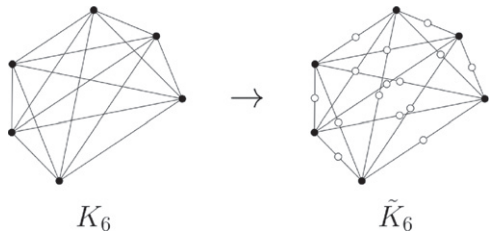

Figure 1. An example of a complete graph, K_6, and its corresponding dotted-complete graph $\tilde{K}_6$.

(iii) There are two communication channels: one quantum and one classical.

The completeness and soundness conditions are identical to the IP conditions.

We are also interested in interactive protocols that use more than one prover. There are only two differences, first that the verifier can interact with multiple provers instead of just one, and second that the provers are not allowed to communicate. The conditions for completeness and soundness remain unchanged. The analogous complexity class that involves multiple provers is called *multi-prover interactive proof system* and denoted MIP [28]. It is defined as the set of all languages which admit an interactive proof system with one or more non-communicating provers. If the number of provers is fixed to be k, the corresponding complexity class is MIP$[k]$. A closely related class is MIP* where the multiple non-communicating provers share entangled states.

In all of these cases, the verifier is essentially *delegating* a difficult computation to the prover(s). This computation can be universal with respect to the computation model of the prover(s). In our case, this means universal for polynomial-time quantum computations. The number of classical messages exchanged between the verifier and the prover, throughout the run of the protocol, as a function of the input size is known as the *round complexity* of the protocol.

1.1.2. Quantum protocols

Throughout this subsection, we assume the reader is familiar with the teleportation-based and more generally measurement-based quantum computing (MBQC) models, described in detail in [29, 30].

We first summarize the FK protocol [1] which is a QPIPprotocol. It is also known as unconditionally secure, verifiable, universal blind quantum computing. The protocol is 'blind' which means that no information about the computation is leaked to the prover, apart from its size. This property can be exploited by allowing the verifier to insert hidden 'traps' within the computation. The traps are deterministic tests which the verifier can perform in order to verify that the prover is not deviating from the protocol. Blindness ensures that the traps are indistinguishable from the computation.

The basic idea of this protocol is that the verifier prepares and sends qubits to the prover. The prover entangles these qubits and then performs adaptive measurements (sending the measurement outcomes to the verifier) that will overall implement a certain unitary operation, as in the MBQC model of computation. The traps are single isolated qubits, disentangled from the rest of the computation, and when measured in suitable bases give deterministic outcomes that are known to the verifier (but not to the prover). Since the prover is completely blind and does not know which qubits are traps and which are part of the actual computation, any attempt to cheat has some probability to affect the trap and thus be detected.

The FK protocol is based on a universal resource state for the MBQC model, known as the *dotted-complete graph state*. The details of this resource state are not crucial for understanding this paper, apart from the fact that, as part of the FK protocol, the appropriate operators are performed by the untrusted server to prepare this generic state. In particular, a series of controlled-Z operators are performed by the server, according to the dotted-complete graph structure, for entangling the individual qubits prepared in advance by the verifier. These initial qubits, that are sometimes referred to as the input of the FK protocol, are sent to the server at the first stage of the protocol. This fact is used to prove some basic properties needed for our main robustness result, see theorem 2. Therefore, for the purpose of completeness, we state here the definition of the dotted-complete graph state, taken from [1], see also figure 1.

Definition 2 [1]. Let K_N denote the complete graph of N vertices. Define the *dotted-complete graph*, denoted as $\tilde{K}_N$, to be a graph where every edge in K_N is replaced with a new vertex connected to the two vertices originally joined by that edge. We call the quantum state corresponding to $\tilde{K}_N$ the *dotted-complete graph state*. This multi-partite entangled state is prepared by replacing every vertex with a qubit in the state $|+\rangle$ and applying a controlled-Z operator for every edge in the graph.

The family of dotted-complete graph states is universal for quantum computation. Moreover, any other graph state could be obtained from a large enough dotted-complete graph state by applying the appropriate Pauli measurements over some of the vertices (the ones shown in white, in figure 1). Concretely, in order to construct any desired graph of N vertices from a dotted-complete graph $\tilde{K}_N$, Pauli Y measurements are performed in order to keep a specific edge, and Pauli Z measurements in order to remove it (alternatively, one could use the states $|0\rangle, |1\rangle$ for the edges which should be removed, instead of performing a Pauli Z measurement). This can be done blindly in order to hide the target graph. The detailed construction is not important for the rest of this paper and hence omitted (see section 5 in [1]). We give a brief description of FK, shown here as protocol 1.

Protocol 1. Fitzsimons, Kashefi QPIP protocol from [1].

Assumptions

The verifier wants to delegate a quantum computation described by the graph G and specific measurement angles $\phi_i \in \{0, \pi/4 \ldots 7\pi/4\}$ chosen to define a desired computation. He chooses a *security parameter d* and encodes this graph into the topological error correcting scheme of Raussendorf *et al* [31], that can correct or detect errors with weight less than d. The encoded graph, $\mathcal{G}$, will have N qubits. He then considers a random hidden partitioning of the vertices of the dotted-complete graph $\widetilde{\mathcal{K}}_{3N}$ into three subgraphs: the computation graph $\mathcal{G}$, and two different types of isolated traps which the verifier will use in order to test the prover's honesty in performing measurements. All measurements are performed in the XY-plane.

Protocol

(1) Verifier prepares and sends the $3N\,(3\,N\,+\,1)/2$ qubits to the prover (the number of vertices in $\widetilde{\mathcal{K}}_{3N}$). These consist of dummy qubits which are either $|0\rangle$ or $|1\rangle$ (isolating their neighbouring vertices from the rest of the computation) and computation or trap qubits which are of the form $|+_\theta\rangle = \frac{1}{\sqrt{2}}(|0\rangle + e^{i\theta}|1\rangle)$, where $\theta \in \{0, \pi/4 \ldots 7\pi/4\}$.

(2) The prover entangles the qubits according to the structure of the $\widetilde{\mathcal{K}}_{3N}$ graph by applying controlled-Z operations between any pair of vertices that are connected with an edge.

(3) For each qubit i, the verifier computes the measurement angle $\delta_i = \theta_i + \phi_i' + r_i\pi$, where ϕ_i' is the adapted version of the computation angle ϕ_i, and r_i is a randomly chosen bit $\{0, 1\}$. Adapted computation angles are used to account for corrections from previous measurements. The measurement angles ϕ_i, for the trap qubits, are randomly fixed to be 0 or π. However, due to blindness provided by the initial θ_i rotations in the preparation of individual qubits (step 1 above), the value of δ_i is uniformly distributed over the set $\{0, \pi/4 \ldots 7\pi/4\}$. The verifier sends these measurement angles one by one to the prover. The prover measures each corresponding qubit in the $|+_{\delta_i}\rangle, |-_{\delta_i}\rangle$ basis, and sends his reply b_i to the verifier.

(4) The verifier accepts if for all trap qubits, the reported measurement outcome b_t is the same as the expected outcome r_t.

According to definition 1 the quantum channel between verifier and prover is one-way (from the verifier to the prover). Moreover, the constant c, representing the number of qubits that the verifier can possess at any given time, is exactly one. We refer the reader to [1] for a more in depth description of the protocol and its associated concepts. However, we recall the key properties of the protocol in the following lemma.

Lemma 1. *Assuming the verifier wants to delegate the computation of a circuit of size N, the FK protocol has* $O(N^2)$ *round complexity and uses* $O(N^2)$ *qubits with completeness being exactly 1 while the soundness is upper bounded by* $(2/3)^{\lceil \frac{2d}{5} \rceil}$, *where d is the security parameter.*

Proof. It is clear from protocol 1 that the total number of qubits used in the protocol is $3N\,(3\,N\,+\,1)/2$, where N is the number of qubits for the encoded graph $\mathcal{G}$. Additionally, we have the same number of rounds of classical communication, corresponding to the measurements of qubits in the dotted-complete graph (each measurement requires three classical bits to specify the measurement angle and 1 bit to specify the outcome). Both the overall round complexity and the required quantum resources are thus $O(N^2)$.

As described in protocol 1, the verifier accepts if and only if all trap measurements succeed. This is always the case if the prover is honest and follows the instructions, since the trap measurements are deterministic. Therefore, the probability that the verifier accepts when the prover is honest is exactly 1 (completeness). On the other hand, it is shown in [1] that in the case of classical output, the protocol is $(2/3)^{\lceil \frac{2d}{5} \rceil}$-verifiable, meaning the soundness is upper bounded by $(2/3)^{\lceil \frac{2d}{5} \rceil}$. □

Next we summarize the RUV protocol [2] which is a MIP* protocol. It relies on the rigidity of CHSH games [20, 32] to test the honesty of the provers (see traps at the FK protocol), and on gate teleportation to perform the computation. In particular, the verifier directs the provers to perform a series of local measurements of their parts of the shared entangled states. The purpose of this is to check for statistical violations of Bell's inequality. At the same time, the verifier makes the provers teleport quantum states into gates in order to perform his desired quantum computation [20]. Importantly, the verifier alternates between these strategies in such a way that the

provers are not aware (they are blind) in which strategy their measurement belongs. Moreover, the two provers cannot use previous results in order to deviate from the protocol (i.e. there is no adaptive cheating strategy). This is summarized as protocol 2. Due to the rigidity of the CHSH games, as proved in [2], the verifier can determine if the two provers are being honest or not from the statistical outcomes. To ensure the verification of universal computations, the resource preparation stage of the protocol will prepare multiple copies of the states:

$$\left\{P\,|0\rangle,\ \ (HP)_2\left|\psi^{\star}\right\rangle,\ \ (GY)_2\left|\psi^{\star}\right\rangle,\ \ \mathrm{CNOT}_{2,4}P_2Q_4\left(\left|\psi^{\star}\right\rangle\otimes\left|\psi^{\star}\right\rangle\right):\ P,\,Q\in\{I,\,X,\,Y,\,Z\}\right\}.$$

Here, $|\psi^{\star}\rangle$ denotes the Bell state $\frac{|\,00\rangle+|\,11\rangle}{\sqrt{2}}$ which is shared among the two provers. In fact the provers share multiple copies of $|\psi^{\star}\rangle$, each prover having one qubit from each Bell pair. Without loss of generality, we can assume that prover 1 has the first qubit and prover 2 has the second qubit. The Hadamard, Phase and controlled-Not gates (denoted as $\{H,\ G,\ \mathrm{CNOT}\}$) constitute a universal gate set for quantum computation. The subscript indices indicate on which qubits the gate acts. An arbitrary quantum circuit is thus simulated by repeatedly doing gate teleportations, while keeping the computation blind from the two provers the entire time.

Lemma 2. *Assuming the verifier wants to delegate the computation of a circuit of size n, the round complexity of the RUV protocol is* $O(n^c)$, *where there exists some constant* c, *such that* $c \geqslant 8192$.

Proof. To determine an upper bound for the round complexity we only need to inspect the number of rounds of CHSH games, since the protocol randomly alternates between this and the other three subprotocols. As shown in protocol 2, the verifier plays N sets of CHSH games with the two provers. Each set consists of qn_s games and $n_s \geqslant n^{64}$. Additionally, it is required that $N \geqslant (qn_s)^{\alpha-1}$, where $n^{\alpha/2} \geqslant n^{64}$, so $\alpha \geqslant 128$. These conditions are necessary for the correctness of the state tomography and process tomography subprotocols [20]. We then have that $N \geqslant dn^{8128}$, where d is a constant of the form $d = q^{\alpha-1}$. This is the number of sets of CHSH games and hence the number of required games is lower bounded by n^{8192}. It follows that the number of rounds is $O(n^c)$, where $c \geqslant 8192$. Note that by 'lower bounded' we refer to the case when all of the CHSH statistics are consistent and the verifier does not reject. In the case of inconsistent statistics, the verifier can reject before playing n^{8192} games. □

The subprotocols of the RUV protocol are themselves verification protocols as proved in [20]. The subprotocol that we will use is the state tomography protocol. As part of the RUV protocol, it is used to prepare resource states which are *XZ-determined*, i.e. states that are uniquely determined by their traces against X and Z operators. To compose the RUV with the FK protocol we will use a modified version of the state tomography subprotocol, so that we can prepare all states that are allowed inputs for the FK protocol. We will give the modified protocol in the next section.

Protocol 2. Reichardt, Unger, Vazirani MIP* protocol from [20] (for two provers).

Assumptions

The verifier delegates a quantum circuit of size n to two quantum provers. Let $n_s = n^{\alpha/2} \geqslant n^{64}$, $q = 11$, $n_g = qn_s$, $N \geqslant n_g^{\alpha-1}$ and $\delta = 1/(6n^{\alpha/8})$. The two provers share Nn_g Bell states.

Protocol

The verifier alternates randomly between four subprotocols. He chooses the first three with probability $(1-\delta)/3)$ and the last one with probability δ.

(1) *CHSH games.* The verifier referees N sets of sequential CHSH games, each consisting of n_g games between the provers. He rejects if they win less than:

$$\cos^2(\pi/8)\,Nn_g - \frac{1}{2\sqrt{2}}\sqrt{Nn_g\log(Nn_g)}$$

of the games.

(2) *State tomography.* The verifier chooses $K \in [N]$ uniformly at random and referees $K-1$ sets of CHSH games. He sends the questions from the Kth set to prover 1, while running a state tomography protocol with prover 2. In this protocol prover 2 is asked to prepare q-qubit resource states by measuring his halves of the shared Bell states. This will collapse prover 1's states to the same q-qubit resource states up to corrections. The verifier checks this using the CHSH measurement outcomes from prover 1. These outcomes tomographically determine the states that are being prepared. He rejects if the tomography statistics are inconsistent.

(3) *Process tomography.* The verifier chooses $K \in [N]$ uniformly at random and referees $K-1$ sets of CHSH games. He sends the questions from the Kth set to prover 2, while running a process tomography protocol with prover 1. In this protocol prover 1 is asked to perform Bell measurements on his halves of the shared Bell states. The verifier checks this using the CHSH measurement outcomes from prover 2. He rejects if the tomography statistics are inconsistent.

(4) *Computation.* The verifier chooses $K \in [N]$ uniformly at random and refereed $K-1$ sets of CHSH games. In the Kth game he runs a state tomography protocol with prover 2 and a process tomography protocol with prover 1. The combination of these two achieves computation via gate teleportation.

2. Main results

2.1. Robustness

The first result we prove is that the FK protocol is robust with respect to small variations in the input. Throughout this paper, by 'input' we are referring to the quantum states that the verifier sends to the prover and not the computation input. Without loss of generality we can assume that the desired computation that will be delegated to the server has the fixed classical input $0, \ldots 0$. Dealing with arbitrary classical or quantum input is straightforward, as explained in [1], and makes no difference for our result. Hence, for the rest of this paper we define the *input state* of the FK protocol to be the tensor product of the individual qubits prepared by the verifier, comprising the dotted-complete graph before the prover applies controlled-Z to entangle them (these include the computation, trap and dummy qubits).

The fact that FK is robust means that the protocol's input state can be deviated from its ideal value by some small amount and the protocol will continue to function. In particular, this input state could be the output of some other protocol, provided that this state was close to its ideal value. As we will see in the next subsection, the RUV protocol is capable of such a preparation. We start by formally defining robustness in this context.

Definition 3 (Robustness). A verification protocol with quantum input is robust if, given that the protocol input is ϵ-close in trace distance to the ideal input, in the limit where $\epsilon \to 0$ the completeness and soundness bounds remain unchanged.

Mathematically, if we denote the multi-qubit input state as ρ, and the pure states comprising the ideal input as π_i, where i goes from 1 to the number of qubits, we have that:

$$\| \rho - \bigotimes_i \pi_i \|_{\mathrm{Tr}} \leqslant \epsilon. \tag{1}$$

Note that ρ is of the same dimension as $\bigotimes_i \pi_i$ as it does not contain any ancilla qubits from the environment. Given the definition of robustness, we prove that:

Theorem 1. *The FK protocol is robust and given an input which is ϵ-close to its ideal value, the completeness is lower bounded by $1 - 2\epsilon$ and the soundness bound changes by at most $O(\sqrt{\epsilon})$.*

Because we are tracing out the environment, which could be controlled by an adversary, the security of the protocol, with a deviated input state, needs to be re-established. We highlight this in the following proof sketch of theorem 1:

Proof sketch. We first examine soundness which considers the case of a dishonest prover. Intuitively, when the prover is malevolent, he will try to convince the verifier to accept an incorrect outcome and thus deviate from the correct protocol. However, as shown in [1], no matter how much the prover deviates, the probability for the verifier to accept a wrong outcome is bounded. If the input to the protocol is already deviated from the ideal, one could expect that the soundness bound remains unchanged. The effect of a deviated input could be incorporated in the deviated actions of the prover. This is indeed the case when the input is uncorrelated with any external system and we can express the deviation as a CPTP map (see lemma 7 for detailed proof).

In the general case, however, the deviated input could be correlated with subsystems controlled by adversaries. This deviation could be used by the prover to improve his cheating probability. Mathematically this is manifested by the fact that the prover's action in the presence of initial correlations is not in general a trace preserving map. It can be expressed as a linear combination of a CPTP deviation and an inhomogeneous term which could be either positive or negative as shown in the [33]. In this case, we use the ϵ-closeness of the input state to derive a bound of order $O(\sqrt{\epsilon})$ for the norm of the inhomogeneous term. From linearity, and using the previous argument it follows that in the general case the soundness bound changes by at most $O(\sqrt{\epsilon})$ (see lemma 8 for detailed proof).

In the case of completeness, we are assuming the prover is honest. If we start with an ϵ-close input state, because of the linearity of the operators involved, we will end up with an output state that is $O(\epsilon)$-close to the ideal output (see lemma 9 for detailed proof). □

A similar approach to lemma 7 was used in [34] for defining approximate blindness, and in [27] to prove universal composability for blind quantum computing protocols. However, to our knowledge, these results are not strong enough to cover the requirements for the composition with the RUV protocol. In [34] only the blindness property was examined while verifying the computation was not considered. In [27] they considered

local-verifiability which does not take into account for example, the possibility of correlated attacks such as those that are possible when the two provers have a prearranged correlated strategy.

2.2. Composition

One of our main objectives is to construct a device independent version of the FK protocol. The first step, was to show that FK is robust. This property guarantees that if we have an input state that is only approximately the ideal one, the protocol continues to work. We can now break the task of achieving device independent FK into two parts, which we need to compose sequentially.

(1) *State preparation*—use a device independent protocol to prepare on the prover's side a state which is ϵclose to the FK input.

(2) *Verified delegated computation*—run the FK protocol with the prover that has the ϵ-close input state (since robustness allows this).

The advantage of this technique is that we are free to use any protocol for state preparation as long as we have the guarantee of ϵ-closeness. This is due to our strong robustness result, which shows that FK will work even if the deviation in the prepared state is correlated with the prover's cheating strategy in the delegated computation stage. In this paper, we achieve state preparation using the device-independent state tomography sub-protocol of RUV. This sub-protocol has the ϵ-closeness property that we require, as explained in [20]. The resulting composite protocol will have a better round complexity than the full RUV protocol for the verification of quantum computations. The complexity can be improved further if a more efficient state preparation protocol is used. Recently, in an independent work that simultaneously appeared with our arxiv version, a more efficient scheme for state preparation is proposed that is based on a self-testing approach [23] rather than the rigidity of CHSH games [20].

We first clarify some details of the RUV protocol, which are essential in understanding how our composite protocol will work. RUV uses the rigidity property of CHSH games to determine that the provers share multiple copies of the Bell state $|\Phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$, which is XZ-determined. They can then use XZ state tomography to verify the preparation of any other XZ-determined state. In particular, they use it to tomographically verify the preparation of a set of states which can be used to perform universal computation. They also describe how it is possible to extend the protocol in order to have full tomography with the Y operator as well [20]. However, because they are using the $|\Phi^+\rangle$ Bell state, it is only possible to fix the Y operator up to a sign change. That is, the provers can always choose to measure in either the Y or $-Y$ bases without being detected (this corresponds to complex conjugating the states with respect to their representations in the computational basis). In fact this problem has been noticed by others as well [12, 35]. As explained in [20], it is possible to force the provers to consistently choose either Y or $-Y$ for their measurements. This makes the resulting state prepared by state tomography close to either the ideal state or the complex conjugate of the ideal state.

At first glance it would seem that this could be problematic for the FK protocol. We would have to show that running the FK protocol with an input state that is close to the complex conjugated version of the ideal input would be detected by the verifier. Intuitively this is the case, since trap qubits are in the XY-plane and complex conjugating them would lead to different measurement outcomes. We will not prove this and instead provide a simpler solution.

The problem stems from the fact that we are using the XZ-determined $|\Phi^+\rangle$ state. Let us instead consider the state $|\Psi^+\rangle = (|01\rangle + |10\rangle)/\sqrt{2}$. Using theorem 2 from [20], and the fact that $|\Psi^+\rangle$ has stabilizer generator set $\{X \otimes X,\ Y \otimes Y\}$ which belongs to $\{I,\ X,\ Y\}^{\otimes 2}$ we have that this state is XY-determined.

Theorem 2 [20]. *A stabilizer state is determined by any of its sets of stabilizer generators.*

In principle it is possible to run a form of the RUV protocol in which we choose the CHSH games such that we rigidly determine that the provers share multiple copies of the Bell state $|\Psi^+\rangle$ instead of $|\Phi^+\rangle$. Analogous to the previous case, the extended form of the protocol would then fix the Z operator up to a sign change (instead of the Y operator). This means that the provers can always perform a reflection about the XY plane with no noticeable changes. However, the XY plane states are invariant under such a reflection. We can therefore use this to prepare the input which will be used by the FK protocol. The only problem we encounter is that we also require the preparation of $|0\rangle$ and $|1\rangle$ states which act as dummy qubits in the FK protocol [1]. As described in protocol 1 these dummy qubits are measured in order to 'break' the dotted-complete graph into the computation graph and the two trap graphs. The problem is that the XY plane reflection has the effect of flipping the computational basis states (state $|0\rangle$ becomes $|1\rangle$ and state $|1\rangle$ becomes $|0\rangle$). However this deviation

(flip) has to be applied globally otherwise it affects the statistics of the CHSH game and thus the verifier rejects [20]. Such a global flip is detected by the FK protocol. A formal proof is given in lemma 10, section 4, while below we give a sketch of the proof.

In the honest scenario for the FK protocol, the measurement of a dummy qubit in state $|1\rangle$ introduces an additional Z correction to its neighbouring qubits (this is because we are using the controlled-Z operation for entangling qubits). Hence in a malicious setting the effect that a flip has on a trap qubit with an odd number of neighbouring dummy qubits, leads to an extra Z operation. Such a Z flip changes a $|+_\theta\rangle$ state to $|-_\theta\rangle$. Thus, the measurement of this trap qubit will deterministically fail and the verifier will detect this. On the other hand, since the verifier chooses the input, he can always pad the computation such that the overall graph has trap qubits with an odd number of neighbour dummy qubits. This is due to the fact that in a dotted-complete graphs (definition 2), some of the traps will have $N-1$ neighbouring dummy qubits. Therefore, if the size of the input computation N is odd, the verifier need only pad the computation size to become $N+1$.

Now we are in a position to construct the composite protocol which composes RUV with FK. We give a modified version of the state tomography protocol of RUV (see protocol 3). Proof that protocol 3 is valid verification protocol is given in section 4. The purpose of this modification is to verifiably prepare the minimal resource states which are subsequently used as inputs for the FK protocol.

Protocol 3. Modified state tomography protocol.

Assumptions

Let $S=\{(1,0,0),(0,1,0),(0,0,1),\frac{1}{\sqrt{2}}(1,1,0),\frac{1}{\sqrt{2}}(1,-1,0),\frac{1}{\sqrt{2}}(1,0,1),\frac{1}{\sqrt{2}}(1,0,-1),\frac{1}{\sqrt{2}}(0,1,1),\frac{1}{\sqrt{2}}(0,1,-1)\}$.

Let M_v be a 2 outcome projective measurement defined by the projectors: $\frac{1}{2}(I+\vec{v}\cdot(X,Y,Z))$ and $\frac{1}{2}(I-\vec{v}\cdot(X,Y,Z))$.

Let the tuple $(\vec{a},\vec{b})\in S\times S$ denote the measurements M_a for prover 1 and M_b for prover 2 that they need to perform on their halves on an entangled state when instructed by the verifier. Sets of such tuples define CHSH games. For example the set $\{(1,0,0),(0,0,1)\}\times\{\frac{1}{\sqrt{2}}(1,0,1),\frac{1}{\sqrt{2}}(1,0,-1)\}$ defines the XZ CHSH game. Given S, there are six such sets of CHSH games (two XZ, two XY and two YZ) [20]. For a suitable numbering of these games, let $CHSH_i$ be the ith CHSH game, $i\in\{1,\ldots 6\}$.

Protocol

The verifier alternates uniformly at random between the following subprotocols:

(1) *CHSH games.* Verifier referees $6N$ sets of sequential CHSH games, such that each group of N sets is one of the six possible CHSH types of games. Each set consists of n_g games between prover 1 and prover 2. For each group of N CHSH games the verifier rejects if the two provers win less than:

$$\cos^2(\pi/8)Nn_g-\frac{1}{2\sqrt{2}}\sqrt{Nn_g\log(Nn_g)}$$

of the games.

(2) *State tomography.* Verifier chooses $K\in[N]$ uniformly at random and also randomly chooses CHSH_i as one of the six possible CHSH games. Then he referees $K-1$ sets of CHSH_i games, sending the questions from the Kth set to prover 1, while running a state tomography protocol with prover 2. In this protocol prover 2 is asked to prepare resource states by measuring his halves of the shared Bell states. This will collapse prover 1's states to the same resource states up to corrections. In the context of composition, these resource states will constitute the FK input. The verifier uses the measurement outcomes of prover 1 to tomographically check this preparation. He rejects if the tomography statistics are inconsistent. In the end, if the verifier accepts, he concludes that with high probability prover 1 has a state which is close in trace distance to the tensor product of resource states. The formal statement of this fact is given in theorem 5, taken from [20], and more precisely in equation (42) from lemma 12.

The composite protocol, given as protocol 4, is the sequential composition of the modified state tomography of protocol 3 with both provers followed by the FK protocol with prover 1. Note that since prover 1 is involved in both state tomography as well as the FK protocol, the strong version of the robustness property is required. This is to address the effect of any potential correlated attacks where provers 1 and 2 have agreed in advance on a strategy. The deviations of prover 2, in the preparation stage, could be correlated with the deviations of prover 1 during the computation stage (FK). This is the first rigorous proof of a protocol that involves lifting the FK protocol to the entangled provers setting. Additionally note that prover 2 is only required to perform quantum measurements and therefore can be viewed as an untrusted measurement device. Hence our protocol is a device-independent single-server verification protocol unlike RUV which is a multi-server protocol. We give here the correctness and soundness of this protocol and show that it is more efficient than the RUV protocol (theorem 3) while in section 4 we give the proof of this theorem.

Protocol 4. Composite verification protocol.

(1) Run the modified state tomography protocol (protocol 3).

(2) From the states prepared by this protocol on prover 1's side, select the input for FK and run the FK protocol with prover 1. (Protocol 1.)

Theorem 3. *Assuming the verifier wants to delegate the computation of a quantum circuit of size n, protocol 4 is a MIP* verification protocol having completeness lower bounded by* $1 - O(n^{-1/128})$*, soundness upper bounded by* $(\frac{2}{3})^{\lceil \frac{2d}{5} \rceil} + O(n^{-1/12})$*, where* d *is the security parameter of the FK protocol, and round complexity* $O(n^c)$*, where there exists some constant* c *such that* $c > 2048$.

While the obtained round complexity is an improvement over RUV (lemma 2) it is still far from practical. However, we believe our approach serves as a proof of principle, that this type of composition can be beneficial. It also highlights where improvements could be made. It is the state tomography subprotocol that increases the round complexity, while the FK protocol has a relatively low complexity[4]. The detailed proofs are given in section 4.

2.3. Fault tolerance

In constructing our composite verification protocol, we used the robustness of the FK protocol. Our last result is to characterize the difference between robustness and fault tolerance and to show that the FK protocol can be made fault tolerant using a topological error correcting code. Note that in [1] a quantum error correcting code is used to further bound the deviation in the prover's cheating strategy and thus boost the soundness parameter of the protocol. However, in that case traps were not encoded and so in the presence of noise and faulty devices, the acceptance probability, even for an honest prover, would decrease considerably, as we will show. We remedy this problem in our fault tolerant version of the FK protocol. Consequently, our composite protocol can also be made fault tolerant provided that the state tomography part is run on top of an error correcting code.

As mentioned before robustness is a protocol's ability to continue to function given a deviated input. Fault tolerance is when a protocol functions correctly in the presence of error prone devices. The essential assumption for robustness is that the actual (multi-qubit) input is ϵ-close to its ideal value. Fault tolerant protocols, on the other hand, assume that errors can occur at each individual qubit. The faulty devices are usually represented by the action of a partially depolarizing channel: $\mathcal{E} = (1 - p)[I] + \frac{p}{3}([X] + [Y] + [Z])$. Here p is the probability of error, and the square brackets indicate the action of an operator. This leads to the following observation:

Lemma 3. *Let* $\sigma = \otimes_{i=1}^{n} \rho_i$ *be a system of n qubits. Assume each qubit goes through a partially depolarizing channel* $\mathcal{E}$ *having probability of error* $p > 0$*. Let the state of the system, after all qubits have passed through the channel, be* $\sigma' = \otimes_{i=1}^{n} \mathcal{E}(\rho_i)$*. We have that* $\| \sigma - \sigma' \|_{\mathrm{Tr}} \leqslant min(1, np)$ *and there exist states* σ *for which* $\| \sigma - \sigma' \|_{\mathrm{Tr}} = 1$.

This means that the deviation of an n-qubit system from the ideal input is not bounded by some constant amount. This is intuitively clear, since by adding more qubits, we introduce more errors and the state of the composite system is further from its intended value. In contrast to this, when considering robustness, the distance between the actual and ideal state is bounded by an arbitrarily small quantity. We will now address how can we do verification in an error prone setting.

Lemma 4. *Assume we run the FK protocol with* N_T *traps and each qubit is subject to the action of a partially depolarizing channel* $\mathcal{E}$ *having probability of error* $p > 0$*. Given the simplifying assumption that if a qubit is changed (through the action of an X, Y or Z operator) it will produce an incorrect measurement outcome, the completeness of this protocol is upper bounded by* $(1 - p)^{N_T}$.

Protocol 5. Fault tolerant FK protocol.

Assumptions

The verifier wants to compute the execution of a measurement graph G having n qubits. Both the verifier and prover's devices are subject to noise modelled as a partially depolarizing channel acting on the preparation of the qubits and the application of the quantum gates. For single qubits the channel is described by:

$$\mathcal{E}_1 = (1 - p)[I] + \frac{p}{3}([X] + [Y] + [Z]). \tag{2}$$

And for two qubit states by:

$$\mathcal{E}_2 = (1 - p)[I \otimes I] + \frac{p}{15}([I \otimes X] + \cdots + [Z \otimes Z]). \tag{3}$$

Additionally assume $p \leqslant p_{\mathrm{correct}}$, where p_{correct} is a threshold such that depolarizing noise bellow this threshold is corrected by the topologically protected code from [31].

[4] Note that the round complexity of FK could be further reduced to linear, if one is willing to admit a higher upper bound for soundness.

(Continued.)

Let $\mathcal{G}^{\nu}$ denote a *brickwork state* encoding the graph G and containing one trap qubit, as explained in [1]. Let $\mathcal{L}^{\nu}$ denote a fault tolerant encoding of the graph $\mathcal{G}^{\nu}$ using the topologically protected code from [31]. The encoding is done as explained in [36], hence $\mathcal{L}^{\nu}$ will be decorated lattices (see figures 1, 2 in [36]). The index ν denotes the randomness in the θ angles for the encoding as chosen by the verifier. Let $S^{\tilde{\nu}} = \mathcal{L}^{\nu_1} \otimes \mathcal{L}^{\nu_2} \otimes \cdots \otimes \mathcal{L}^{\nu_N}$, where $R/\log\left(\frac{cn}{cn-1}\right) < N < R/\log\left(\frac{cn}{cn-1}\right) + O(1)$, for some constants $R > 1$, $c > 2$ and $\tilde{\nu} = \{\nu_1 \cdots \nu_N\}$. We will refer to $S^{\tilde{\nu}}$ as a *sequence* of encodings.

Protocol

(1) The verifier chooses $R > 1$ and constructs the random set $\tilde{\nu}$.

(2) The verifier prepares the qubits for the sequence $S^{\tilde{\nu}}$ and sends them to the prover along with instructions on how to construct $S^{\tilde{\nu}}$.

(3) The verifier sends measurement instructions to the prover in order to compute the executions of the encoded graphs.

(4) The prover sends the measurement outcomes to the verifier.

(5) Steps 3 and 4 repeat until the verifier either accepts or rejects.

The verifier rejects if any of the traps fail. He takes the outcome of the computation to be the majority outcome over all computations (graphs $\mathcal{L}^{\nu_i}$).

It is evident that assuming faulty devices where each qubit behaves as if it crossed a partially depolarizing channel, the completeness of the protocol becomes exponentially small (as function of the number of traps). This is clearly unsatisfactory. The arguably simplest solution would be to alter the acceptance condition of the protocol. Since it is unlikely that all trap measurements succeed, even for honest prover, the verifier should accept a result if the traps that succeed are above some fixed fraction.

Lemma 5. *Assume we run a modified FK protocol with N_T traps and each qubit is subject to the action of a partially depolarizing channel $\mathcal{E}$ having probability of error $p > 0$. The modification is that the verifier accepts if there are fewer than $N_T(p + \epsilon)$ mistakes at trap measurements, where $\epsilon > 0$ is a suitably chosen small number. The completeness of this protocol is lower bounded by* $1 - \exp(-2\epsilon^2 N_T)$.

The above modification resolves the issue raised regarding the completeness bound. However, if we were to make such modification, we have the following consequence for the soundness of the protocol:

Lemma 6. *Assume we run a modified FK protocol with N_T traps,* N *qubits in total and each qubit is subject to the action of a partially depolarizing channel $\mathcal{E}$ having probability of error $p > 0$. The modification is that the verifier accepts if there are fewer than $N_T(p + \epsilon)$ mistakes at trap measurements, where $\epsilon > 0$ is a suitably chosen small number. The soundness of this protocol is upper bounded by* $\binom{N_T}{N}\left(\frac{2}{3}\right)^{\left\lceil \frac{2d}{5} \right\rceil}$.

We can see that introducing a threshold of acceptance leads to an increased bound on soundness. Again, expected, since we allow the prover to tamper with some of the traps (and, by extension, with the computation as well) without rejecting the output. To solve these problems we need to use a fault tolerant code. The FK protocol already uses a fault tolerant code to encode the computation graph. However this is done in order to boost the value of the soundness parameter. The trap qubits are not encoded with the code (only the computation is). Thus we propose a modified FK protocol. This is described in protocol 5. In this protocol we encode both computations and traps in a fault tolerant code and use sequential repetitions (also used in [17]). This leads to our final main result:

Theorem 4. *Under the assumption of a faulty setting where qubit preparation and quantum gates are subject to partially depolarizing noise having bounded probability p, protocol 5 is a valid verification protocol having completeness 1, soundness upper bounded by* $(1/2)^R$, *where R is a constant such that* $R > 1$. *The protocol has round complexity* $O(n^2)$.

The proof of this theorem (and previous lemmas) are given in section 5. An important point to make is that the composite protocol we constructed can also be made fault tolerant. To achieve this the state tomography protocol should be run on top of a fault tolerant code. As mentioned in [20], in principle, this is straightforward for blind, verified computation, since the provers can work on top of a quantum error-correcting code and entanglement can be distilled with the help of the verifier [37, 38].

3. Proof of robustness

In this section we prove the robustness of the FK protocol. We start by first proving a simpler result, namely the robustness of the protocol under the assumption that the input is uncorrelated with any external system. We then remove this assumption and use our results to prove the main theorem, necessary for the composition with the RUV protocol.

Lemma 7. *If the initial input state of the FK protocol is* ϵ*-close to the ideal input state and uncorrelated with any external system, the soundness bound does not change.*

Proof. We will follow the same proof technique as in [1] and show that the soundness bound does not change. This is done by incorporating the assumption of a deviated input into that proof. The outcome density operator of the protocol is denoted $B_j(\nu)$, where ν denotes the verifier's choices of input variables and j ranges over the prover's choices of possible actions ($j = 0$ is the correct/honest action). If the outcome is incorrect it means that all of the traps have passed, but the computation is not correct. This is associated with the following projection operator [1]:

$$P_{\text{incorrect}} = \left(\mathbb{I} - \left|\Psi_{\text{ideal}}\right\rangle\left\langle\Psi_{\text{ideal}}\right|\right) \bigotimes_{t\in T} \left|\eta_t^{\nu_T}\right\rangle\left\langle\eta_t^{\nu_T}\right|. \tag{4}$$

Here, $|\Psi_{\text{ideal}}\rangle\langle\Psi_{\text{ideal}}|$ is the ideal output state, and $\otimes_{t\in T}|\eta_t^{\nu_T}\rangle\langle\eta_t^{\nu_T}|$ is the state associated with the trap qubits. Notice that we are projecting to a state in which the output is orthogonal to its ideal value, and the traps are correct. This expresses the fact that the verifier will accept an incorrect computation. The associated probability for that event is $p_{\text{incorrect}}$ and can be expressed as:

$$p_{\text{incorrect}} = \sum_{\nu} p(\nu)\, Tr\left(P^{\nu}_{\text{incorrect}} B_j(\nu)\right). \tag{5}$$

Which is a weighted average of the incorrect outcome probabilities (expressed by the trace operator) over all possible input states. The outcome density operator can be written as:

$$B_j(\nu) = \text{Tr}_P\left(\sum_b \left|b + c_r\right\rangle\langle b|\, C_{\nu_C,b}\Omega P\overbrace{\left(\underbrace{\left(\otimes^P |0\rangle\langle 0|\right)}_{\text{Prover's qubits}} \otimes \underbrace{\left|\Psi^{\nu,b}\right\rangle\left\langle\Psi^{\nu,b}\right|}_{\text{Input state}}\right)}^{\text{Joint system state}\sigma^{\nu,b}}\right.$$
$$\left. P^{\dagger}\Omega^{\dagger}C^{\dagger}_{\nu_C,b}\,|b\rangle\left\langle b + c_r\right|\right). \tag{6}$$

Notice the following, as explained in [1]:

- we are tracing over the prover's qubits;
- we have denoted the joint state, comprised of the input and the prover's qubits, as $\sigma^{\nu,b}$;
- j ranges over the prover's possible strategies ($j = 0$ is the honest strategy);
- b indicates the possible branches of computation parametrized by the measurement results sent by the prover to the verifier;
- c_r indicates corrections that need to be performed on the final, classical output due to the MBQC computation together with the random phase introduced by the verifier;
- P is the computation that we want the prover to do;
- Ω is the prover's deviation from the desired computation;
- $C_{\nu_C,b}$ are the corrections the prover applies to its quantum output depending on the measurement outcomes (as in the measurement-based model);

we now need to incorporate the approximate input state into this operator. We will not use the ϵ-closeness of the deviated state to the ideal one, and prove a stronger result, that the soundness bound does not change *regardless* of the input state. Concretely, assume the deviated input is:

$$\rho^{\nu,b} = \mathcal{E}\left(\left|\Psi^{\nu,b}\right\rangle\left\langle\Psi^{\nu,b}\right|\right), \tag{7}$$

where $\mathcal{E}$ is a CPTP map which represents any deviation from the ideal input state either from incorrect preparation, a malicious prover or faulty devices. This is equivalent to applying some unitary U to the input state tensored with some environment qubits that are traced out. We can express this mathematically as:

$$\rho^{\nu,b} = \mathrm{Tr}_E\left(U\left(\left(\otimes^E |0\rangle\langle 0|\right)\otimes\left|\Psi^{\nu,b}\right\rangle\left\langle\Psi^{\nu,b}\right|\right)U^\dagger\right). \tag{8}$$

The joint system state $\sigma^{\nu,b}$ becomes[5]:

$$\sigma^{\nu,b} = \mathrm{Tr}_E\left(\left(\otimes^P |0\rangle\langle 0|\right)\otimes U\left(\left(\otimes^E |0\rangle\langle 0|\right)\otimes\left|\Psi^{\nu,b}\right\rangle\left\langle\Psi^{\nu,b}\right|\right)U^\dagger\right). \tag{9}$$

Let us consider a new unitary $V = (\mathbb{I}\otimes U)$. This allows us to rewrite the joint system state as:

$$\sigma^{\nu,b} = \mathrm{Tr}_E\left(V\left(\left(\otimes^{E+P} |0\rangle\langle 0|\right)\otimes\left|\Psi^{\nu,b}\right\rangle\left\langle\Psi^{\nu,b}\right|\right)V^\dagger\right). \tag{10}$$

Since P, the computation, is a unitary operator, there must exists some unitary V' such that $V = P^\dagger V' P$. Substituting this into the previous expression gives us:

$$\sigma^{\nu,b} = \mathrm{Tr}_E\left(P^\dagger V' P\left(\left(\otimes^{E+P} |0\rangle\langle 0|\right)\otimes\left|\Psi^{\nu,b}\right\rangle\left\langle\Psi^{\nu,b}\right|\right)P^\dagger V'^\dagger P\right). \tag{11}$$

Incorporating equation (11) into the expression for $B_j(\nu)$ in equation (6) we obtain:

$$\begin{aligned} B_j(\nu) = \mathrm{Tr}_P\Bigg(&\sum_b \left|b + c_r\right\rangle\langle b|\, C_{\nu_C,b}\Omega P \\ &\left(\mathrm{Tr}_E\left(P^\dagger V' P\left(\left(\otimes^{E+P} |0\rangle\langle 0|\right)\otimes\left|\Psi^{\nu,b}\right\rangle\left\langle\Psi^{\nu,b}\right|\right)P^\dagger V'^\dagger P\right)\right) \\ &P^\dagger\Omega^\dagger C^\dagger_{\nu_C,b}\,|b\rangle\left\langle b + c_r\right|\Bigg). \end{aligned} \tag{12}$$

The assumption of the lemma is that the input state is not correlated with any external system. Hence, the spaces E and P are independent. This means that the prover's deviation, Ω, is not acting on E and therefore we can 'push' the inner trace operator to the beginning of the equation. Also using the fact that $PP^\dagger = P^\dagger P = \mathbb{I}$, we obtain:

$$\begin{aligned} B_j(\nu) = \mathrm{Tr}_{P+E}\Bigg(&\sum_b \left|b + c_r\right\rangle\langle b| \\ &C_{\nu_C,b}\Omega V' P\left(\left(\otimes^{P+E} |0\rangle\langle 0|\right)\otimes\left|\Psi^{\nu,b}\right\rangle\left\langle\Psi^{\nu,b}\right|\right)P^\dagger V'^\dagger\Omega^\dagger C^\dagger_{\nu_C,b} \\ &|b\rangle\left\langle b + c_r\right|\Bigg). \end{aligned} \tag{13}$$

We can now include the input deviation given by V' into the prover deviation Ω, by considering $\Omega' = \Omega V'$. This is possible because we are bounding the probability over all possible deviations, Ω, of the prover in the computation and all possible deviations, V', from the preparation part. Thus, we can consider this to be a single, global, deviation given by Ω'.

$$\begin{aligned} B_j(\nu) = \mathrm{Tr}_{P+E}\Bigg(&\sum_b \left|b + c_r\right\rangle\langle b| \\ &C_{\nu_C,b}\Omega' P\left(\left(\otimes^{P+E} |0\rangle\langle 0|\right)\otimes\left|\Psi^{\nu,b}\right\rangle\left\langle\Psi^{\nu,b}\right|\right)P^\dagger\Omega'^\dagger C^\dagger_{\nu_C,b} \\ &|b\rangle\left\langle b + c_r\right|\Bigg). \end{aligned} \tag{14}$$

As a result, the above equation has the same form as the undeviated input scenario of equation (6). This makes sense since all we have done is to incorporate the deviation of the input into the prover's cheating strategy. The original proof continues as it is in [1], and the bound remains unchanged

$$p_{\text{incorrect}} \leqslant \left(\frac{2}{3}\right)^{\left\lceil\frac{2d}{5}\right\rceil}. \tag{15}$$

□

[5] Here and in the following expressions we have used the fact that the partial trace is linear and can therefore be moved outside.

The type of robustness guaranteed by this lemma is not sufficient to prove the security of any protocol that composes RUV with FK. For example if we use prover 2 of RUV to prepare the input of the FK protocol for prover 1, this input is in general correlated with prover 2's system. To address this issue, we use from [20] the following corollary of the gentle measurement lemma and the special Kraus representation in the presence of initial correlations given in [33].

Corollary 1 [20]. *Let* ρ *be a state on* $\mathcal{H}_1 \otimes \mathcal{H}_2$, *and let* π *be a pure state on* $\mathcal{H}_1$. *If for some* $\delta \geqslant 0$, $\mathrm{Tr}(\pi\ \mathrm{Tr}_2\rho) \geqslant 1 - \delta$, *then*

$$||\rho - \pi \otimes \mathrm{Tr}_1\rho||_{\mathrm{Tr}} \leqslant 2\sqrt{\delta} + \delta. \tag{16}$$

Lemma 8. *If the initial input state of the FK protocol is* ϵ*-close to the ideal input state the soundness bound changes by at most* $O(\sqrt{\epsilon})$.

Proof. Consider a composite correlated state ρ_{AB} where systems A and B are not communicating and let $\rho_A = \mathrm{Tr}_B(\rho_{AB})$ and $\rho_B = \mathrm{Tr}_A(\rho_{AB})$. If ρ_A is used as input for the FK protocol, the existence of correlations (not present in the previous lemma) can be exploited by an adversarial prover. Hence the deviation can no longer be expressed as a CPTP map over this subsystem. As it is shown in [33], in presence of initial correlations defined as:

$$\rho_{\mathrm{corr}} = \rho_{AB} - \rho_A \otimes \rho_B \tag{17}$$

the evolution of the subsystem ρ_A is the following:

$$\rho_A \rightarrow \mathcal{E}(\rho_A) + \delta\rho_A. \tag{18}$$

Here $\mathcal{E}$ is a CPTP map and $\delta\rho_A$ is an inhomogeneous term which is added to the CPTP evolution due to the presence of initial correlations. In addition we have the following property:

$$\delta\rho_A = \mathrm{Tr}_B\left(U_{AB}\rho_{\mathrm{corr}}U_{AB}^{\dagger}\right). \tag{19}$$

We can see that substituting ρ_A in the outcome density operator of the FK protocol gives different soundness bound than the one in lemma 7. The difference stems from the extra $\delta\rho$ term. However we can use the fact that ρ is ϵ-close to the ideal state (lemma assumption) to show that the norm of $\delta\rho_A$ is at most of order $O(\sqrt{\epsilon})$. To prove this we first find a bound for the norm of ρ_{corr}, and since $\delta\rho_A$ is just a CPTP map applied to ρ_{corr} it follows that the norm of $\delta\rho_A$ has the same bound. Moreover, the action of the FK protocol can be modelled as a CPTP map, therefore acting on $\delta\rho_A$ will not increase the norm. It follows that the overall soundness bound changes by at most $O(\sqrt{\epsilon})$.

If we denote the ideal state as $|\psi\rangle$, we know that:

$$||\rho_A - |\psi\rangle\langle\psi|||_{\mathrm{Tr}} \leqslant \epsilon. \tag{20}$$

It is also known, from the relationship between fidelity and trace distance, that:

$$1 - \langle\psi|\,\rho_A\,|\psi\rangle \leqslant ||\rho_A - |\psi\rangle\langle\psi|||_{\mathrm{Tr}}. \tag{21}$$

Combining these two yields:

$$\langle\psi|\,\rho_A\,|\psi\rangle \geqslant 1 - \epsilon. \tag{22}$$

Recall that $\mathrm{Tr}(|\psi\rangle\langle\psi|\,\rho_A) = \langle\psi|\,\rho_A\,|\psi\rangle$, using equation (22) and corollary 1 (where ρ is substituted with ρ_{AB} and π with $|\psi\rangle\langle\psi|$) we have:

$$||\rho_{AB} - |\psi\rangle\langle\psi| \otimes \rho_B||_{\mathrm{Tr}} \leqslant 2\sqrt{\epsilon} + \epsilon. \tag{23}$$

The trace norm of ρ_{corr} is simply the trace distance between ρ_{AB} and $\rho_A \otimes \rho_B$ as can be seen from the definition. Using the triangle inequality, we have:

$$||\rho_{AB} - \rho_A \otimes \rho_B||_{\mathrm{Tr}} \leqslant ||\rho_{AB} - |\psi\rangle\langle\psi| \otimes \rho_B||_{\mathrm{Tr}} + |||\psi\rangle\langle\psi| \otimes \rho_B - \rho_A \otimes \rho_B||_{\mathrm{Tr}}. \tag{24}$$

For the last term, using the additivity of trace distance with respect to tensor product, we get:

$$|||\psi\rangle\langle\psi| \otimes \rho_B - \rho_A \otimes \rho_B||_{\mathrm{Tr}} \leqslant |||\psi\rangle\langle\psi| - \rho_A||_{\mathrm{Tr}} + ||\rho_B - \rho_B||_{\mathrm{Tr}} = \epsilon. \tag{25}$$

Combining these last three inequalities we obtain:

$$||\rho_{AB} - \rho_A \otimes \rho_B||_{\mathrm{Tr}} \leqslant 2\sqrt{\epsilon} + 2\epsilon. \tag{26}$$

Since $0 \leqslant \epsilon \leqslant 1$, the bound is of order $O(\sqrt{\epsilon})$. We have therefore bounded the norm of ρ_{corr} and thus the norm of $\delta\rho_A$.

We can now take our expression for the deviated input from equation (18) and substitute it into equation (8), from lemma 7. Since trace is a linear operation, it will result in the addition of an inhomogeneous term to each equation that involves the outcome density operator. But since the inhomogeneous term has bounded trace norm, and the action of the outcome density operator is trace preserving, it follows that we obtain the same bound as in lemma 7 with the addition of an extra term of order $O(\sqrt{\epsilon})$. This concludes the proof. □

Lemma 9. *If the initial input state of the FK protocol is* ϵ*-close to the ideal input state, the completeness is lower bounded by* $1 - 2\epsilon$.

Proof. In the simplest sense, the FK protocol can be abstractly thought of as a CPTP map $\mathcal{P}$, that takes some input state to an output state. Since we are assuming the prover is honest, the output state will be $B_0(\nu)$. However, this is in the case where the input is assumed to be ideal. We are dealing with a deviated input, hence our output state will be $B_0'(\nu)$. Writing these out explicitly we have:

$$B_0(\nu) = \mathcal{P}\left(\left|\Psi^\nu\right\rangle\left\langle\Psi^\nu\right|\right), \tag{27}$$

$$B_0'(\nu) = \mathcal{P}\left(\rho^\nu\right), \tag{28}$$

where, ρ^ν is the deviated input, and by assumption

$$\| \rho^\nu - \left|\Psi^\nu\right\rangle\left\langle\Psi^\nu\right| \|_{\mathrm{Tr}} \leqslant \epsilon. \tag{29}$$

Note that in the following we do not need to consider non CPTP map evolution since the provers are assumed to behave honestly. Hence even in the presence of initial correlation, the subsystem will evolve according to the desired CPTP map of the protocol. However CPTP maps cannot increase the trace distance, which leads to:

$$\| B_0(\nu) - B_0'(\nu) \|_{\mathrm{Tr}} \leqslant \| \left|\Psi^\nu\right\rangle\left\langle\Psi^\nu\right| - \rho^\nu \|_{\mathrm{Tr}} \leqslant \epsilon. \tag{30}$$

This also applies for projection operators and if in particular we consider P_{correct}, the projection onto the correct output state, we also have that:

$$\| P_{\mathrm{correct}} B_0(\nu) - P_{\mathrm{correct}} B_0'(\nu) \|_{\mathrm{Tr}} \leqslant \| B_0(\nu) - B_0'(\nu) \|_{\mathrm{Tr}} \leqslant \epsilon. \tag{31}$$

Next we use the reverse triangle inequality, which gives us:

$$\left| \| P_{\mathrm{correct}} B_0(\nu) \|_{\mathrm{Tr}} - \| P_{\mathrm{correct}} B_0'(\nu) \|_{\mathrm{Tr}} \right| \leqslant \| P_{\mathrm{correct}} B_0(\nu) - P_{\mathrm{correct}} B_0'(\nu) \|_{\mathrm{Tr}} \leqslant \epsilon. \tag{32}$$

And since we are dealing with positive definite operators, we know that:

$$\| P_{\mathrm{correct}} B_0(\nu) \|_{\mathrm{Tr}} = \frac{1}{2} \mathrm{Tr}\left(P_{\mathrm{correct}} B_0(\nu)\right), \tag{33}$$

$$\| P_{\mathrm{correct}} B_0'(\nu) \|_{\mathrm{Tr}} = \frac{1}{2} \mathrm{Tr}\left(P_{\mathrm{correct}} B_0'(\nu)\right). \tag{34}$$

But $\mathrm{Tr}(P_{\mathrm{correct}} B_0(\nu)) = 1$ (the completeness when we have ideal input), so:

$$\left| 1 - \mathrm{Tr}\left(P_{\mathrm{correct}} B_0'(\nu)\right) \right| \leqslant 2\epsilon. \tag{35}$$

Lastly, because $Tr(P_{\mathrm{correct}} B_0'(\nu)) \leqslant 1$, we get:

$$1 - 2\epsilon \leqslant Tr\left(P_{\mathrm{correct}} B_0'(\nu)\right). \tag{36}$$

Thus, the probability of accepting a correct outcome, under the assumption that the input state is ϵ-close to the ideal input, is greater than $1 - 2\epsilon$. □

It is now easy to see that the proof of theorem 1 follows directly from definition 3 and lemmas 8 and 9. Having the robustness property, the FK protocol can receive an input, which is ϵ-close to its ideal value, from another protocol. As we have shown, even if this input is correlated with an external system, we can still perform the verification as long as we have ϵ-closeness.

4. Proof of compositionality

To prove the security of the composite protocol (theorem 3), we first need to prove that the FK protocol rejects with high probability a state close to a reflection about the XY-plane (lemma 10). Then we prove that the

modified state tomography protocol (protocol 3), satisfies the ϵ-closeness property required by the (robust) FK. This is achieved by showing lemmas 11 and 12. Finally we give the proof of theorem 3.

Lemma 10. *If the initial input state of the FK protocol is* ϵ*-close to a reflection about the* XY*-plane of the ideal input state, the protocol will reject it with high probability.*

Proof. First we note that the input to the FK protocol consists of XY-plane states and dummy qubits which are either $|0\rangle$ or $|1\rangle$. The XY-plane states are invariant under the reflection, while the dummy states will be flipped. Assume that there is a trap that has an odd number of (dummy) neighbours. The verifier knows that he sent the state $|+_\theta\rangle$ and expects to make a Z correction if the number of $|1\rangle$ neighbours is odd. However, if instead of what the verifier expects, there is an overall reflection with respect to the XY-plane, then for each of the neighbours of the trap there will be a new Z correction (the $|0\rangle$ will become $|1\rangle$ inducing a Z, while the $|1\rangle$ will become $|0\rangle$ undoing the previous Z correction, which is equivalent with another Z correction since $Z^2 = \mathbf{1}$). Therefore, if the neighbours of a trap are odd in number, he will expect the exact opposite result and will deterministically detect the deviation. For this to happen it suffices that the verifier makes sure that at least one trap has odd number of neighbours, something that can be easily achieved. Therefore, the FK protocol will always reject the reflected ideal input state. Given that the input is ϵ-close to this, we have shown in lemma 8 that the outcome density operator changes by at most $O(\sqrt{\epsilon})$ from its ideal value. Thus, the output state is $O(\sqrt{\epsilon})$ close to the reflected ideal input state. It follows that the protocol will reject this state with at most probability $1 - O(\sqrt{\epsilon})$. □

In proving the correctness of our protocol we first need to show the correctness of the modified state tomography protocol. Here we focus on the main results that we use for showing the correctness and security of this protocol. We start with a theorem from [20]:

Theorem 5 [20]. *Fix* $\mathcal{Q} = \{\pi^1, \ldots, \pi^{2^q}\}$ *a complete, orthonormal set of q-qubit XZ-determined pure states. For a sufficiently large constant* α *and for sufficiently large n, let* $m = m(n) \geqslant qn$ *and* $N \geqslant m^{\alpha-1}$. *Let* $\sigma \in [m]^{qn}$ *be a list of distinct indices. Consider a combination of the following two protocols between the verifier, Eve, and the provers, Alice and Bob.*

(1) CHSH games: in the first protocol, Eve referees Nm sequential CHSH games. She accepts if

$$\left|\left\{j \in [Nm]\colon A_j B_j = X_j \oplus Y_j\right\}\right| \geqslant \cos^2(\pi/8)Nm - \tfrac{1}{2\sqrt{2}}\sqrt{Nm\log(Nm)}\,. \tag{37}$$

(2) State tomography: in the second protocol, Eve chooses $K \in [N]$ *uniformly at random. She referees* $(K-1)m$ *CHSH games. For the Kth set, she referees a state tomography protocol with parameters q, n, m,* $\mathcal{Q}$ *and* σ. *She accepts if the following criteria are satisfied:*

$$\max_{o \in \left[2^q\right]} \left|\#\left\{j\colon O_j = o\right\} - n/2^q\right| \leqslant 4^q\sqrt{n\log n} \tag{38}$$

$$\max_{o \in \left[2^q\right], P \in \{I,X,Z\}^{\otimes q}} \left|\tau^{o,P} - \mathrm{Tr}\left(\pi^o P\right)\right| \leqslant 4^q\sqrt{(\log n)/n}\,. \tag{39}$$

The combined protocol satisfies the following completeness and soundness conditions:

Completeness: if Alice and Bob use Nm shared EPR states to play the CHSH games according to an ideal strategy, and if Bob uses an ideal strategy with respect to the projections $\mathcal{Q}$ *on the Kth set of m EPR states in the state tomography protocol, then in both protocols*

$$\Pr[\text{Eve accepts}] \geqslant 1 - O\left(n^{-1/2}\right). \tag{39}$$

Soundness: assume that for both protocols, $\Pr(\text{Eve accepts}) \geqslant 1 - n^{-1/3}$. *Let* ρ *be Alice's state in the second protocol after* $(K-1)m$ *games and conditioned on Bob's messages* $O_1, \ldots, O_n$. *Then there exists an isometry* $\mathcal{X}^A\colon \mathcal{H}_A \hookrightarrow (\mathbb{C}^2)^{\otimes m} \otimes \mathcal{H}'_A$ *such that letting* $\rho_{\sigma,j}$ *be* $\mathcal{X}^A \rho \mathcal{X}^{A\dagger}$ *reduced to Alice's qubits* $\{\sigma(j, i)\colon i \in [q]\}$,

$$\Pr\left[\left|\left\{j\in[n]:\ \mathrm{Tr}\left(\rho_{\sigma,j}\pi^{O_j}\right)\geqslant 1-O\left(n^{-1/16}\right)\right\}\right|\geqslant\left(1-O\left(n^{-1/16}\right)\right)n\right]\geqslant 1-4n^{-1/12}. \quad (40)$$

Here, the probability is over K, the first $(K-1)m$ *games and* $O_1, \ldots, O_n$.

We give the following corollary to this theorem:

Corollary 2. *By changing the measurement operators accordingly, a state tomography protocol for q-qubit* XY*-determined (YZ-determined) states exists, and achieves the same completeness and soundness bound as the one from theorem 5.*

Proof. As mentioned, if we consider the extended CHSH game (comprising of six CHSH games) and try to rigidly determine the existence of a tensor product of $|\Psi^+\rangle$ states, we can fix the strategies of the provers up to an XY plane reflection. In particular, the results of an XZ state tomography in the original setting hold here for XY (YZ) tomography. Therefore, it is possible to certify the preparation of q-qubit XY-determined (YZ-determined) states. □

We can now present the main lemma proving that protocol 3 is a verification protocol.

Lemma 11. *Protocol 3 has completeness lower bounded by* $1-O(n^{-1/2})$ *and soundness upper bounded by* $O(n^{-1/12})$.

Proof. According to corollary 2, the six state tomography protocols that constitute protocol 3 are valid verification protocols achieving the same bounds for completeness and soundness as the original protocol from theorem 5. We will ignore the case of XY plane reflections since, as we have shown in lemma 10, these are detected with overwhelming probability by the FK protocol. These protocols can be 'stitched' together in the same way the subprotocols of the RUV protocol are stitched together. In fact, our case requires a much simpler analysis since the six state tomography protocols are independent of each other. This means that in each subprotocol, the verifier is not basing his questions on the results of any previous subprotocol. This nonadaptive technique contrasts the RUV protocol in which the questions were adaptive. In the case of honest provers, the verifier accepts if all subprotocols succeed. For each one, we know from theorem 5 that the probability of acceptance is $\geqslant 1-O(n^{-1/2})$, hence for the whole protocol the probability of acceptance is $\geqslant(1-O(n^{-1/2}))^6=1-O(n^{-1/2})$. Thus, we see that the completeness bound remains unchanged. For soundness, assuming the provers are dishonest we know, again from theorem 5, that the probability of accepting an incorrect outcome is $\leqslant 4n^{-1/12}$. In our protocol, the provers can be dishonest in any of the six subprotocols, therefore, by a union bound the probability of accepting an incorrect outcome is $\leqslant 6\cdot 4n^{-1/12}=24n^{-1/12}$. Therefore, we can say that the soundness of our protocol is upper bounded by $O(n^{-1/12})$. □

We are now able to give the proof of our main result (theorem 3) which concerns the properties of the composite protocol (protocol 4). To do this, we require an additional property.

Lemma 12. *Assume the verifier wants to prepare a state* ρ *consisting of tensor products of qubits which are all determined in either the XZ, XY or YZ bases. A successful run of protocol 3 certifies that, prover 1 has a state* ρ' *such that* ρ *and* ρ' *are close in trace distance.*

Proof. The proof is partially given in [20]. In the state tomography protocol, a prover prepares multiple copies of a resource state. In [20] it is stated that if the verifier accepts, then, with high probability, a subset of states of the prover are close in trace distance to copies of the resource state.

The soundness condition of theorem 5 states that:

$$\Pr\left[\left|\left\{j\in[n]:\ \mathrm{Tr}\left(\rho_{\sigma,j}\pi^{O_j}\right)\geqslant 1-O\left(n^{-1/16}\right)\right\}\right|\geqslant\left(1-O\left(n^{-1/16}\right)\right)n\right]\geqslant 1-4n^{-1/12}. \quad (41)$$

It is shown in [20] that this condition translates to the fact that with probability at least $1-O(n^{-1/48})$ we have:

$$\|\,\rho_S\left(O_{1,n}\right)-\otimes_{j\in S}\,\pi^{O_j}\,\|_{\mathrm{Tr}}\leqslant O\left(n^{-1/64}\right) \quad (42)$$

Where S is uniformly random subset of size $O(n^{1/64})$. If we denote $O(n^{-1/64})$ as ϵ, $O(n^{-1/48})$ as p, $\rho_S(O_{1,n})$ as ρ_ϵ and $\bigotimes_{j\in S}\pi^{O_j}$ as ρ_{id} then the state ρ', that prover 1 has, is:

$$\rho' = (1-p)\rho_\epsilon + p\left(I-\rho_\epsilon\right). \tag{43}$$

We can see that, for sufficiently large n, the values of p and ϵ tend to 0. Consequently, ρ' approaches ρ_ϵ and ρ_ϵ approaches the ideal state, ρ_{id}. Computing the trace distance between ρ' and ρ_{id}, we obtain:

$$\| \rho' - \rho_{id} \|_{\mathrm{Tr}} = \|(1-p)\rho_\epsilon + p\left(I-\rho_\epsilon\right) - \rho_{id}\|_{\mathrm{Tr}} \leqslant \| \rho_\epsilon - \rho_{id} \|_{\mathrm{Tr}} + \| p\left(I-2\rho_\epsilon\right)\|_{\mathrm{Tr}}. \tag{44}$$

And using inequality (42), we have:

$$\| \rho' - \rho_{id} \|_{\mathrm{Tr}} \leqslant O\left(n^{-1/64}\right) + 2p = O\left(n^{-1/64}\right) + O\left(n^{-1/48}\right) = O\left(n^{-1/64}\right). \tag{45}$$

Therefore, the state that prover 1 has, conditioned on his messages $O_{1,n}$, is close to the state comprised of copies of the resource states. Depending on which type of state tomography is done, the resource states are determined in either the XZ, XY or YZ bases. □

Proof of theorem 3. According to lemmas 11 and 12, protocol 3 is capable of preparing with high probability, a multi-qubit state ρ on prover 1's side, such that ρ is ϵ-close to a tensor product of states determined in either the XZ, XY or YZ bases. In fact, each subprotocol is capable of such a preparation. For the two XY state tomography protocols we choose the resource state to be:

$$|+\rangle \otimes \left|+_{\pi/4}\right\rangle \otimes \left|+_{2\pi/4}\right\rangle \otimes \left|+_{3\pi/4}\right\rangle \otimes \left|+_{4\pi/4}\right\rangle \otimes \left|+_{5\pi/4}\right\rangle \otimes \left|+_{6\pi/4}\right\rangle \otimes \left|+_{7\pi/4}\right\rangle. \tag{46}$$

For the two XZ state tomography protocols we choose the resource state to be:

$$|0\rangle \otimes |1\rangle \otimes |0\rangle \otimes |1\rangle. \tag{47}$$

This allows us to prepare multi-qubit states on prover 1's side which are close in trace distance to the FK input consisting of XY-plane states and dummy qubits (the $|0\rangle$, $|1\rangle$ qubits). If we denote as ρ_1 the multi-qubit state consisting of multiple copies of the XY resource state and ρ_2 as the multi-qubit state consisting of multiple copies of the XZ resource state, then the FK input is effectively $\rho_1 \otimes \rho_2$. Lemma 12 shows that with high probability prover 1 will have a state ρ_1' that is ϵ_{prep}-close to ρ_1 and a state ρ_2' that is ϵ_{prep}-close to ρ_2, where $\epsilon_{\mathrm{prep}} = O(n^{-1/64})$. Therefore, $\rho_1' \otimes \rho_2'$ is $2\epsilon_{\mathrm{prep}}$-close to $\rho_1 \otimes \rho_2$. Moreover, in [20] it is proven in the state tomography protocol prover 1 is completely blind regarding his state. Given this, and using theorem 1, we can compose the modified state tomography protocol (protocol 3) with the FK protocol to achieve a new blind verification protocol. The state $\rho_1' \otimes \rho_2'$ is used as input for the FK protocol, since it is ϵ-close to the ideal input, where $\epsilon = 2\epsilon_{\mathrm{prep}}$.

The bound on completeness for the new protocol can be computed from the completeness bounds of the constituent protocols. In the honest provers setting, the verifier's acceptance probability for modified state tomography is $1 - O(n^{-1/2})$, and for FK with deviated input it is $1 - O(\sqrt{\epsilon}) = 1 - O(n^{-1/128})$. Multiplying these together and taking the leading order terms we find that completeness of the protocol is upper bounded by $1 - O(n^{-1/128})$.

For soundness, in the dishonest setting if the verifier would reject in either modified state tomography or FK then he would reject in the new protocol as well. The bound on soundness for modified state tomography is $O(n^{-1/12})$ and for FK is $(\frac{2}{3})^{\lceil \frac{2d}{5} \rceil}$, where d is the security parameter of the FK protocol that specifies the size of the encoding for the computation graph. From a union bound we get that the soundness for our composite approach is $(\frac{2}{3})^{\lceil \frac{2d}{5} \rceil} + O(n^{-1/12})$.

The last part of the proof deals with the round complexity of our composite approach. In the previous proof of lemma 12 we mentioned that prover 1's state restricted to subset of $O(n^{1/64})$ qubits is close in trace distance to the ideal state. However we need n to be sufficiently large so that this subset of qubits can encompass the entire FK input. We know that the FK input comprises of $O(|\mathcal{C}|^2)$ qubits, where $\mathcal{C}$ is the computation the verifier wants to perform. This means, that we need $O(|\mathcal{C}|^{128})$ qubits in total so that we can claim that a state of $O(|\mathcal{C}|^2)$ qubits is close to its intended value. Recall that from thereom 5, the number of rounds for state tomography is $O(n^\alpha)$, where we know from [20] that $\alpha > 16$. This means that the total number of rounds must be $O(|\mathcal{C}|^c)$, where $c > 128 \cdot 16 = 2048$. If we relabel n to be $|\mathcal{C}|$ then the round complexity is $O(n^c)$. □

5. Proof of fault tolerance

The main result of this section is the proof of theorem 4 that gives a fault tolerant FK protocol. We first prove lemmas 3–6 that as stressed in section 2.3, highlights why we cannot use results similar to the robustness and why the simplest approaches fail. Then we proceed in the proof of theorem 4.

Proof of lemma 3. We can compute a bound on the trace distance between an arbitrary qubit ρ_i and $\mathcal{E}(\rho_i)$:

$$\| \rho_i - \mathcal{E}(\rho_i)\|_{\mathrm{Tr}} = \| \rho_i - (1-p)\rho_i - (p/3)([X]+[Y]+[Z])\rho_i([X]+[Y]+[Z])\|_{\mathrm{Tr}}, \tag{48}$$

$$\| \rho_i - \mathcal{E}(\rho_i)\|_{\mathrm{Tr}} = p \| \rho_i - (1/3)([X]+[Y]+[Z])\rho_i([X]+[Y]+[Z])\|_{\mathrm{Tr}}. \tag{49}$$

But we know that the trace distance is upper bounded by 1, so:

$$\| \rho_i - \mathcal{E}(\rho_i)\|_{\mathrm{Tr}} \leqslant p. \tag{50}$$

Now we compute the trace distance between $\sigma = \otimes_{i=1}^{n} \rho_i$ and $\sigma' = \otimes_{i=1}^{n} \mathcal{E}(\rho_i)$:

$$\| \sigma - \sigma' \|_{\mathrm{Tr}} = \| \otimes_{i=1}^{n} \rho_i - \otimes_{i=1}^{n} \mathcal{E}(\rho_i)\|_{\mathrm{Tr}} \leqslant \sum_{i=1}^{n} \| \rho_i - \mathcal{E}(\rho_i)\|_{\mathrm{Tr}}. \leqslant np \tag{51}$$

Since the trace distance is upper bounded by 1 and since np can exceed 1 for sufficiently large n, we have:

$$\| \sigma - \sigma' \|_{\mathrm{Tr}} \leqslant \min(1, np). \tag{52}$$

Consider $\sigma = \otimes_{i=1}^{n} |0\rangle\langle 0|$. Under the action of the depolarizing channel the trace distance between σ and σ' is $n \, \| |0\rangle\langle 0| - \mathcal{E}(|0\rangle\langle 0|)\|_{\mathrm{Tr}}$. However $\| |0\rangle\langle 0| - \mathcal{E}(|0\rangle\langle 0|)\|_{\mathrm{Tr}} = \sqrt{\frac{2p}{3}}$, therefore the distance between σ and σ' is $n\sqrt{\frac{2p}{3}}$. For sufficiently large n this can clearly reach the maximum value of 1. □

Proof of lemma 4. Because of the action of the partially depolarizing channel, each trap qubit has a probability p of being changed. We make the simplifying assumption that an affected qubit will produce a wrong measurement result. This assumption is only valid for completeness, where we assume that the devices are honest but faulty[6]. We then have that the probability of a trap measurement producing a correct outcome is upper bounded by $1-p$. Given that trap measurements are independent of each other, and assuming we have N_T traps, the probability that all trap measurements produce correct outcomes is upper bounded by $(1-p)^{N_T}$. Since the verifier accepts if and only if all trap measurements succeed it follows that the completeness is upper bounded by $(1-p)^{N_T}$. □

Proof of lemma 5. Define the following Bernoulli random variable:

$$X_t = \begin{cases} 1, & \text{if measurement of trap } t \text{ fails,} \\ 0, & \text{otherwise.} \end{cases} \tag{53}$$

Under the simplifying assumption of the previous lemma, we have $\Pr(X_t = 1) = p \geqslant 0$. Next, we define:

$$F = \sum_{t=1}^{N_T} X_t. \tag{54}$$

It is clear that $E(F) = N_T p$. Additionally, using a Hoeffding inequality, we have that:

$$\Pr\left(F \geqslant (p+\epsilon)N_T\right) \leqslant \exp\left(-2\epsilon^2 N_T\right). \tag{55}$$

This gives the probability that the number of failed traps is greater than our threshold of pN_T. The complement of this is the completeness, which is therefore bounded by $1 - \exp(-2\epsilon^2 N_T)$ □

[6] If the devices were dishonest, we would need to take into account the deviation on the trap qubits resulting from malevolent behaviour.

Proof of lemma 6. Recall that soundness is the probability of accepting an incorrect outcome. In the original FK protocol this meant that all the traps succeeded but the computation output was orthogonal to the correct output. This is expressed with the projector $P_\perp \otimes P_T$. Here $P_\perp$ projects the computation output onto the orthogonal state and P_T projects the trap outputs onto the correct outputs. If the accepting condition is given by a threshold of correct traps, the projector must change accordingly. This means that there should not be only one trap projector but one for each accepting situation. Taking the threshold to be pN_T means that the verifier accepts if $T = N_T - pN_T$ traps succeeded. Since these traps can be any combination of T out of the possible N_T, there are $\binom{N_T}{T}$ possible accepting situations. Therefore, the trap projector P_T becomes a sum of $\binom{N_T}{T}$ projectors (one for each accepting choice of traps). It therefore follows from linearity that the soundness bound becomes $\binom{N_T}{T}\left(\frac{2}{3}\right)^{\left\lceil \frac{2d}{5} \right\rceil}$.

□

Proof of theorem 4. In [36], Morimae and Fujii show how a blind quantum computation can be made fault tolerant by encoding it in a topologically protected error-correcting code [31]. The encoding then uses a decoration trick so that the prover only needs to perform XY-plane measurements and this can be done blindly using the UBQC protocol in [1]. Here, we use the same idea to encode a computation which also contains an isolated trap. This follows from the first verification protocol introduced in [1] which uses a brickwork state to perform the computation. Encoding this in the fault tolerant code give us the lattice $\mathcal{L}^\nu$, which according to [36] can be executed blindly by the prover.

Throughout the run of the protocol, if the prover is always honest then the fault tolerant code will correct for any errors (since we have assumed the error rate is smaller than the threshold of correctable errors). This proves that the completeness of the protocol is 1.

To compute the soundness, note that the computed bound for the brickwork state protocol in [1] is:

$$p_{\text{incorrect}} < \left(1 - \frac{1}{2n}\right), \tag{56}$$

where n is the number of qubits in the brickwork state. Similar to the robustness proof, the proof of this bound assumes that the outcome density operator of the protocol is projected onto a state where the trap succeeded but the computation outcome is incorrect. It is can been shown that the same bound as the non-fault tolerant case holds. This means that we have:

$$p_{\text{incorrect}} < \left(1 - \frac{1}{2n'}\right), \tag{57}$$

where n' is the number of qubits in a lattice $\mathcal{L}^{\nu_i}$, out of the N lattices used in the protocol. We note that n' is of the same order as n [31], and we can choose a constant $c > 2$ such that $2n' = cn$. In protocol 5 the verifier creates *independent* encodings $\mathcal{L}^\nu$, each depending on classical randomness. He accepts the sequence of encodings if all trap measurements succeed in each encoding. This means that the prover can deceive the verifier if he can deviate the computation in each encoding $\mathcal{L}^\nu$ while at the same time passing all the traps. However, for any given encoding we know that the probability of this happening is given by $p_{\text{incorrect}}$, and because of independence, the prover will succeed for the sequence with probability:

$$p^N_{\text{incorrect}} < \left(1 - \frac{1}{cn}\right)^N. \tag{58}$$

We know that $N/R > 1/\log\left(\frac{cn}{cn-1}\right)$. However, this is equivalent to:

$$N/R > -1/\log\left(1 - \frac{1}{cn}\right), \tag{59}$$

$$(N/R)\log\left(1 - \frac{1}{cn}\right) < -1. \tag{60}$$

Note that we used the fact that $\log\left(1 - \frac{1}{cn}\right) < 0$. Through exponentiation we get:

$$\left(1 - \frac{1}{cn}\right)^{N/R} < \frac{1}{2}. \tag{61}$$

And we finally obtain:

$$p^N_{\text{incorrect}} < \frac{1}{2^R}. \tag{62}$$

Hence, the probability that the prover deceives the verifier is less than $(1/2)^R$ and so the soundness of the protocol is upper bounded by this value.

Lastly we compute the round complexity of this protocol. For the given sequence we have N encodings and for each encoding we have $O(n)$ qubits[7] and a corresponding round complexity of $O(n)$ to compute the execution of that encoding. It follows that the overall complexity is $O(Nn)$. But we know that $N < R/\log\left(1 + \frac{1}{cn-1}\right) + O(1)$, and given that R is a constant, we can show that N is $O(n)$. This follows from the observation that dividing the function $1/\log\left(1 + \frac{1}{cn-1}\right)$ with $(cn-1)$ gives a constant in the $n \to \infty$ limit:

$$\lim_{n\to\infty} (cn-1)\log\left(1 + \frac{1}{cn-1}\right) = \frac{1}{\ln 2} \tag{63}$$

Incorporating this result yields overall complexity $O(n^2)$. Note that this proof technique works for the case of classical output since we are interested in the classical output of each encoding. The encodings are independent from each other, which allows us to bound the probability for the whole sequence. □

6. Conclusion

We have shown that the single server universal verifiable blind quantum computing protocol of [1] is robust even against general adversaries. This protocol is currently the optimal protocol in terms of the verifier's requirements. The robustness result further strengthens the scheme for realistic applications where the effect of noisy devices should also be considered, as highlighted in a recent experimental demonstration of the protocol [17]. Moreover, it enables us to compose the FK protocol with other quantum verification protocols, extend it to the entangled servers setting and make it device independent. The key property that we proved, is that the protocol remains secure even against correlated attacks. To achieve this, we considered the deviation of the evolution of a correlated subsystem from the evolution of uncorrelated subsystems. The former could be written mathematically as a non-CPTP map which differs from a CPTP map by an inhomogeneous term. However, for inputs which are ϵ-close to the ideal FK input, we showed that this deviation (the inhomogeneous term) is bounded by a term of order $O(\sqrt{\epsilon})$. Our proof technique is generic and can be potentially applied to other multi-party protocols where sequential composition is required. This result complements the *local*-verifiability proof of [27] which is based on the universal composability framework. The latter, in its current form, is insufficient for composing entanglement-based protocols, such as RUV, with the FK protocol because of the possibility of correlated attacks. Our robustness result, however, leads to a stand alone secure composite verification protocol. Additionaly, the proposed composition scheme could potentially be used to extend the composable framework of [27] to incorporate multiple provers.

Our proposed composite protocol achieves verification with a classical client (device independence) and gives improved round complexity in comparison to the RUV protocol. It uses only the (modified) state tomography part of RUV as input for the FK protocol. The improved round complexity of the composite protocol is still too high to allow for any practical implementation in the near future. However, the reason for this high round complexity is the state tomography subprotocol and therefore, any improvement on how to prepare the FK inputs (e.g. by exploiting the shared entanglement of the provers or using self-testing techniques as in [23]) will directly improve the efficiency of our composite protocol as well.

Finally we outlined how to make our verification protocol fault tolerant. To do so we constructed a fault tolerant version of the FK protocol which is interesting in its own right. This complements the work presented in [36] which addresses the fault tolerance of a (non-verifiable) blind quantum computing protocol. We used the same topological error correcting code as [36] and a sequential repetition scheme in order to correct for faulty devices.

Acknowledgments

Shortly before uploading a prepint on the arxiv, the authors became aware of parallel and independent research by Hajdusek, Perez-Delgado and Fitzsimons, which also addresses device-independent verifiable blind quantum computing and appeared the same day in the arxiv [23]. We would like to thank Vedran Dunjko and Theodoros Kapourniotis for useful discussions. PW gratefully acknowledges partial support from COST Action MP1006.

[7] Note that here, unlike in the composite protocol, we have a linear number of qubits for the protocol. This is because we are not using the dotted-complete version of the FK protocol, but the one using a brickwork state. It is explained in [1] that this version of the protocol requires $O(n)$ qubits.

References

[1] Fitzsimons J F and Kashefi E 2012 Unconditionally verifiable blind computation arXiv:1203.5217
[2] Reichardt B W, Unger R F and Vazirani U 2013 Classical command of quantum systems *Nature* **496** 456–60
[3] Aaronson S and Arkhipov A 2011 The computational complexity of linear optics, *Proc. 43rd Annual ACM Symp. on Theory of Computing, STOC'11* (New York: ACM) pp 333–342
[4] Nielsen M A and Chuang I L 2000 *Quantum Computation and Quantum Information* (Cambridge: Cambridge University Press)
[5] Vazirani U and Vidick T 2014 Fully device-independent quantum key distribution *Phys. Rev. Lett.* **113** 140501
[6] Goldwasser S, Micali S and Rackoff C 1989 The knowledge complexity of interactive proof systems *SIAM J. Comput.* **18** 186–208
[7] Watrous J 1999 Pspace has constant-round quantum interactive proof systems, *Proc. 40th Annual Symp. on Foundations of Computer Science, FOCS'99* (Washington, DC: IEEE Computer Society) p 112
[8] Broadbent A, Fitzsimons J and Kashefi E 2009 Universal blind quantum computation, *Proc. 50th Annual Symp. on Foundations of Computer Science, FOCS 09* (Washington, DC: IEEE Computer Society) pp 517–526
[9] Pappa A, Chailloux A, Wehner S, Diamanti E and Kerenidis I 2012 Multipartite entanglement verification resistant against dishonest parties *Phys. Rev. Lett.* **108** 260502
[10] Aharonov D, Ben-Or M and Eban E 2010 Interactive proofs for quantum computations *Proc. Innovations in Computer Science 2010, ICS2010* p 453
[11] Aaronson S and Arkhipov A 2014 Bosonsampling is far from uniform *Quantum Inf. Comput.* **14** 1383–423
[12] McKague M 2013 Interactive proofs for BQP via self-tested graph states arXiv:1309.5675
[13] Kapourniotis T, Kashefi E and Datta A 2014 Blindness and verification of quantum computation with one pure qubit *9th Conf. on the Theory of Quantum Computation, Communication and Cryptography (TQC 2014)* vol 27 pp 176–204
[14] Dupuis F, Nielsen J and Salvail L 2012 Actively secure two-party evaluation of any quantum operation *Advances in Cryptology CRYPTO 2012* (*Lecture Notes in Computer Science* vol 7417) ed R Safavi-Naini and R Canetti (Berlin: Springer) pp 794–811
[15] Hallgren S, Smith A and Song F 2011 Classical cryptographic protocols in a quantum world *Advances in Cryptology CRYPTO 2011* (*Lecture Notes in Computer Science* vol 6841) ed P Rogaway (Berlin: Springer) pp 411–428
[16] Broadbent A, Gutoski G and Stebila D 2013 Quantum one-time programs *Advances in Cryptology-CRYPTO 2013* pp 344–60
[17] Barz S, Fitzsimons J F, Kashefi E and Walther P 2013 Experimental verification of quantum computation *Nat. Phys.* **9** 727–31
[18] Aharonov D and Vazirani U 2012 Is quantum mechanics falsifiable? A computational perspective on the foundations of quantum mechanics arXiv:1206.3686
[19] Aaronson S 2015 *The Scott Aaronson 25.00$ Prize* Accessed http://scottaaronson.com/blog/?p=284
[20] Reichardt B W, Unger F and Vazirani U 2012 A classical leash for a quantum system: command of quantum systems via rigidity of CHSH games arXiv:1209.0448
[21] Morimae T 2014 Verification for measurement-only blind quantum computing *Phys. Rev.* A **89** 060302
[22] Hayashi M and Morimae T 2015 Verifiable measurement-only blind quantum computing with stabilizer testing arXiv:1505.07535
[23] Hajdušek M, Pérez-Delgado C A and Fitzsimons J F 2015 Device-independent verifiable blind quantum computation arXiv:1502.02563
[24] Unruh D 2010 Universally composable quantum multi-party computation *Proc. 29th Annu. Int. Conf. on Theory and Applications of Cryptographic Techniques EUROCRYPT'10* pp 486–505
[25] Renner R and König R 2005 Universally composable privacy amplification against quantum adversaries *Theory of Cryptography* (*Lecture Notes in Computer Science* vol 3378) ed J Kilian (Berlin: Springer) pp 407–25
[26] Ben-Or M, Horodecki M, Leung D, Mayers D and Oppenheim J 2005 The universal composable security of quantum key distribution *Theory of Cryptography* (*Lecture Notes in Computer Science* vol 3378) ed J Kilian (Berlin: Springer) pp 386–406
[27] Dunjko V, Fitzsimons J F, Portmann C and Renner R 2014 Composable security of delegated quantum computation *Advances in Cryptology ASIACRYPT* (*Lecture Notes in Computer Science* vol 8874) (Berlin: Springer) pp 406–425
[28] Ben-Or M, Goldwasser S, Kilian J and Wigderson A 1988 Multi-prover interactive proofs: how to remove intractability assumptions, *Proc. 12th Annual ACM Symp. on Theory of Computing, STOC 88* (New York: ACM) pp 113–131
[29] Raussendorf R and Briegel H J 2001 A one-way quantum computer *Phys. Rev. Lett.* **86** 5188–91
[30] Danos V, Kashefi E and Panangaden P 2007 The measurement calculus *J. ACM* **54**
[31] Raussendorf R, Harrington J and Goyal K 2007 Topological fault-tolerance in cluster state quantum computation *New J. Phys.* **9** 199
[32] Clauser J F, Horne M A, Shimony A and Holt R A 1969 Proposed experiment to test local hidden-variable theories *Phys. Rev. Lett.* **23** 880–4
[33] Hayashi H, Kimura G and Ota Y 2003 Kraus representation in the presence of initial correlations *Phys. Rev.* A **67** 062109
[34] Dunjko V, Kashefi E and Leverrier A 2011 Universal blind quantum computing with weak coherent pulses arXiv:1108.5571
[35] McKague M and Mosca M 2011 Generalized self-testing and the security of the 6-state protocol, *Proc. 5th Conf. on Theory of Quantum Computation, Communication, and Cryptography, TQC10* (Berlin: Springer) pp 113–130
[36] Morimae T and Fujii K 2012 Blind topological measurement-based quantum computation *Nat. Commun.* **3** 1036
[37] Sheng Y-B and Zhou L 2015 Deterministic entanglement distillation for secure double-server blind quantum computation *Sci. Rep.* **5** 7815
[38] Morimae T and Fujii K 2013 Secure entanglement distillation for double-server blind quantum computation *Phys. Rev. Lett.* **111** 020502